Well-Balanced Schemes for the Initial Boundary Value Problem for 1D Scalar Conservation Laws

M. Nolte   D. Kröner

Abstract

We consider well-balanced numerical schemes for the following 1D scalar conservation law with source term:
$$\partial_t u + \partial_x f(u) + z'(x)\, b(u) = 0.$$
More precisely, we are interested in the numerical approximation of the initial boundary value problem for this equation. While our main concern is a convergence result, we also have to extend Otto's notion of entropy solutions to conservation laws with a source term. To obtain uniqueness, we show that a generalization, the so-called entropy process solution (see [7]), is unique and coincides with the entropy solution.

If the initial and boundary data are in L^∞, we can establish convergence to the entropy solution. Showing that the numerical solutions are bounded we can extract a weak∗-convergent subsequence. Identifying its limit as an entropy process solution requires some effort as we cannot use Kružkov-type entropy pairs here. We restrict ourselves to the Engquist-Osher flux and identify the numerical entropy flux for an arbitrary entropy pair. By the uniqueness result, the scheme then approximates the entropy solution and a result by Vovelle then guarantees that the convergence is strong in L^p, $1 \leq p < \infty$.

1 Introduction

In this paper we will consider well-balanced numerical schemes for the initial boundary value problem

$$
\begin{aligned}
\partial_t u + \partial_x f\left(u\right) + z'\left(x\right) b\left(u\right) &= 0 \quad \text{in } \Omega_T, \\
u\left(\cdot, 0\right) &= u_0 \quad \text{in } \Omega, \\
u\left(x_l, \cdot\right) &= u_l \quad \text{on } \left]0, T\right[, \ \text{if } f'\left(u\right) > 0, \\
u\left(x_r, \cdot\right) &= u_r \quad \text{on } \left]0, T\right[, \ \text{if } f'\left(u\right) < 0,
\end{aligned}
\tag{1}
$$

where $\Omega := \left]x_l, x_r\right[$, $\Omega_T := \Omega \times \left]0, T\right[$ and

$$
\begin{aligned}
u_0 \in L^\infty\left(\Omega\right), \quad u_r, u_l \in L^\infty\left(\left]0, T\right[\right), \quad f \in C^2(\mathbb{R}, \mathbb{R}), \\
z \in H^{1,\infty}\left(\Omega\right), \quad b \in C^1\left(\mathbb{R}\right), \ b' \in L^\infty\left(\mathbb{R}\right)
\end{aligned}
\tag{2}
$$

are given data. Furthermore we define

$$
D\left(s\right) := \int_0^s \frac{f'\left(\xi\right)}{b\left(\xi\right)} \, \mathrm{d}\xi
\tag{3}
$$

and assume that

$$
D \in C^1\left(\mathbb{R}\right), \quad D\left(\mathbb{R}\right) = \mathbb{R}, \quad \inf_{\mathbb{R}} D' > 0.
\tag{4}
$$

For the corresponding initial value problem Greenberg et al. [9] have developed and investigated a well-balanced scheme which is much more efficient than standard schemes. In particular stationary solutions can be approximated by these schemes with less numerical investment than by the classical schemes. Further results including error estimates and numerical experiments for the initial value problem have been obtained by Gosse [8], Perthame [3]. Corresponding results in multiple space dimensions can be found in [2] and [4].

If $b = 0$ there are many results for the initial boundary value problem concerning existence of entropy solutions (see [1], [14]) and convergence of numerical solutions to the entropy solution [15]. Now in this paper we want to analyse the combination of both, i.e. well-balanced schemes for the case $b \neq 0$.

The outline of this paper is as follows. In Section 2 we will repeat the definition of entropy and entropy process solutions. The numerical scheme will be described in Section 3. The main result and its proof will be given in Section 4. Finally in Section 5 we will show the advantages of these well-balanced schemes in some numerical experiments.

2 Entropy Process Solutions

In [14] Otto has defined boundary entropy pairs for problems without source terms, which we will generalize here to problems with source terms. As in [14] we will define boundary entropy pairs as follows.

Definition 2.1 (Boundary Entropy Pairs)   *1. Let $\eta \in C^2(\mathbb{R})$ be convex, $q \in C^1(\mathbb{R})$ and $q' = \eta' f'$. Then (η, q) is called an entropy pair for the partial differential equation in (1).*

2. *Let $H \in C^2(\mathbb{R}^2)$, $Q \in C^1(\mathbb{R}^2)$, $(H(\cdot, w), Q(\cdot, w))$ be an entropy pair for all $w \in \mathbb{R}$ so that*
$$H(w, w) = \partial_1 H(w, w) = Q(w, w) = 0 \quad \forall w \in \mathbb{R}.$$
Then (H, Q) is called a boundary entropy pair in the sense of Otto for the partial differential equation in (1).

3. *Let (η, q) be an entropy in the sense of (1) and $w \in \mathbb{R}$ such that*
$$\eta(w) = \eta'(w) = q(w) = 0.$$
Then (η, q) is called a boundary entropy pair for the partial differential equation in (1).

There is a close and simple relation between the boundary entropy pairs in the sense of (2) and (3) in Definition 2.1.

Remark 2.2   *1. If (H, Q) is a boundary entropy pair in the sense of Otto (see Definition 2.1(2)) and $w \in \mathbb{R}$ then*
$$\eta(u) := H(u, w), \quad q(u) := Q(u, w)$$
form a boundary entropy pair (see Definition 2.1(3)).

2. *If (η, q) is a boundary entropy pair (see Definition 2.1(3)) and $w \in \mathbb{R}$ such that $\eta(w) = \eta'(w) = q(w) = 0$ then*
$$H(u, v) := \eta(u - (v - w)), \quad Q(u, v) := q(u - (v - w))$$
form a boundary entropy pair in the sense of Otto (see Definition 2.1(2)).

Now we use the boundary entropy pairs from Definition 2.1(3) to define an entropy solution of the initial boundary value problem (1).

Definition 2.3 (Entropy Solution) Let $u \in L^\infty(\Omega_T)$, $C := \|u\|_{L^\infty(\Omega_T)}$ and for all boundary entropy pairs (η, q) and all $\varphi \in C_0^\infty(\overline{\Omega} \times [0, T[)$, $\varphi \geq 0$, let

$$\int_0^T \int_\Omega \eta(u) \, \partial_t \varphi + q(u) \, \partial_x \varphi - \eta'(u) \, z'(x) \, b(u) \, \varphi \, \mathrm{dx} \, \mathrm{dt}$$
$$+ \int_\Omega \eta(u_0(x)) \, \varphi(x, 0) \, \mathrm{dx} \tag{5}$$
$$+ \operatorname*{Lip}_{[-C, C]}(f) \int_0^T \eta(u_r(t)) \, \varphi(x_r, t) + \eta(u_l(t)) \, \varphi(x_l, t) \, \mathrm{dt} \geq 0,$$

where

$$\operatorname*{Lip}_{[-C, C]}(f) := \sup_{-C \leq u, v \leq C} \frac{|f(u) - f(v)|}{|u - v|}.$$

Then u is called an entropy solution of (1). If $z' = 0$ this definition corresponds to the definition of entropy solution by Otto (see [12]).

The basic idea of the convergence proof for the numerical solutions consists in replacing the exact solution u in (5) by the numerical solution $u_{\Delta x}$ as defined in Definition 3.1. First we will prove that $u_{\Delta x}$ remains bounded for $\Delta x \to 0$ and then we have to control the limit in (5) for $\Delta x \to 0$. This can be easily done on the basis of the following lemma.

Lemma 2.4 (Eymard, Gallouët, Herbin [7]) Let $(u_k)_{k \in \mathbb{N}} \subset L^\infty(\Omega_T)$ be a bounded sequence. Then there exists a subsequence $(u_{k'})_{k' \in \mathbb{N}}$ and a function $\mu \in L^\infty(\Omega_T \times]0, 1[)$ such that for any $g \in C^0(\mathbb{R})$ we have

$$g\left(u_{k'}\right) \to \int_0^1 g\left(\mu\left(x, t, \alpha\right)\right) \, \mathrm{d}\alpha \quad \text{weak} * .$$

This gives us the motivation for the following definition of the entropy process solution.

Definition 2.5 (Entropy Process Solution, see [7]) Let $\mu = \mu(x, t, \alpha)$ be an $L^\infty(\Omega_T \times]0, 1[)$-function, $C := \|\mu\|_{L^\infty(\Omega_T \times]0,1[)}$ and for all boundary entropy pairs (η, q) and all $\varphi \in C_0^\infty(\overline{\Omega} \times [0, T[)$, $\varphi \geq 0$, let

$$\int_0^T \int_\Omega \int_0^1 \eta\left(\mu\right) \partial_t \varphi + q\left(\mu\right) \partial_x \varphi - \eta'\left(\mu\right) z'\left(x\right) b\left(\mu\right) \varphi \, \mathrm{d}\alpha \, \mathrm{d}x \, \mathrm{d}t$$
$$+ \int_\Omega \eta\left(u_0(x)\right) \varphi(x, 0) \, \mathrm{d}x \tag{6}$$
$$+ \operatorname*{Lip}_{[-C,C]}\left(f\right) \int_0^T \eta\left(u_r\left(t\right)\right) \varphi\left(x_r, t\right) + \eta\left(u_l\left(t\right)\right) \varphi\left(x_l, t\right) \, \mathrm{d}t \geq 0.$$

Then μ is called an entropy process solution of (1).

Now we are going to prove that the entropy process solution is unique. Later in Section 4 we will show that the numerical solutions converge to an entropy process solution. Then the uniqueness will give us the existence of an entropy solution.

Theorem 2.6 (Uniqueness of the Entropy Process Solution) Let $\mu, \nu \in L^\infty(\Omega_T \times]0, 1[)$ be two entropy process solutions of (1) with respect to the inital data $u_0 \in L^\infty(\Omega)$ and to the boundary data $u_l, u_r \in L^\infty(]0, T[)$. Then $\mu = \nu$ a.e. on $\Omega_T \times]0, 1[$ and

$$u\left(x, t\right) := \int_0^1 \mu\left(x, t, \alpha\right) \, \mathrm{d}\alpha$$

is an entropy solution of (1).

Notice that this theorem also proves uniqueness of the entropy solution.

In order to prepare the proof of this theorem we need the following definition and some lemmata.

Definition 2.7 (Semi-Kružkov Entropy Pairs) For $k \in \mathbb{R}$ *define the semi-Kružkov entropy pairs*

$$\eta_k^+ (s) := (s - k)^+, \ q_k^+ (s) := \text{sign}^+ (s - k) (f (s) - f (k))$$
$$\eta_k^- (s) := (s - k)^-, \ q_k^- (s) := \text{sign}^- (s - k) (f (s) - f (k))$$

where $s^+ := \text{sign}^+(s)s$, $s^- := \text{sign}^-(s)s$ *and*

$$\text{sign}^+ (s) = \begin{cases} 1 & \text{for } s > 0, \\ 0 & \text{for } s \leq 0, \end{cases} \quad \text{sign}^- (s) = \begin{cases} 0 & \text{for } s \geq 0, \\ -1 & \text{for } s < 0. \end{cases}$$

It can easily be shown that for any semi-Kružkov entropy pair (η, q) and any $u \in [-C, C]$ we have

$$|q (u)| \leq \underset{[-C,C]}{\text{Lip}} (f) \, \eta (u).$$

Now it turns out that we can use the semi-Kružkov entropy pairs instead of the boundary entropy pairs in (6). This will be made precise in the next lemma.

Lemma 2.8 Let $\mu \in L^\infty(\Omega_T \times]0, 1[)$, $\mu = \mu(x, t, \alpha)$ be an entropy process solution of (1), $C := \|\mu\|_{L^\infty(\Omega_T \times]0,1[)}$, $k \in \mathbb{R}$ and (η_k^+, q_k^+), (η_k^-, q_k^-) be semi-Kružkov entropy pairs in the sense of Definition 2.7. Then we can replace (η, q) in (6) by (η_k^+, q_k^+) and (η_k^-, q_k^-) respectively, i.e. for all $\varphi \in C_0^\infty(\overline{\Omega} \times [0, T[)$, $\varphi \geq 0$, we have

$$\begin{aligned}
\int_0^T \int_\Omega \int_0^1 &\eta_k^\pm (\mu) \, \partial_t \varphi + q_k^\pm (\mu) \, \partial_x \varphi \, \mathrm{d}\alpha \, \mathrm{d}x \, \mathrm{d}t \\
&- \int_0^T \int_\Omega \int_0^1 \text{sign}^\pm (\mu - k) \, z' (x) \, b (\mu) \, \varphi \, \mathrm{d}\alpha \, \mathrm{d}x \, \mathrm{d}t \\
&+ \int_\Omega \eta_k^\pm (u_0 (x)) \, \varphi (x, 0) \, \mathrm{d}x \\
+ \underset{[-C,C]}{\text{Lip}} (f) \int_0^T &\eta_k^\pm (u_r (t)) \, \varphi (x_r, t) + \eta_k^\pm (u_l (t)) \, \varphi (x_l, t) \, \mathrm{d}t \geq 0.
\end{aligned} \tag{7}$$

Proof Let $p_{k,\delta} (x) := -\frac{1}{2\delta^3} (x - k)^4 + \frac{1}{\delta^2} (x - k)^3$ and define

$$\eta_{k,\delta}^+ (x) := \begin{cases} 0 & x \leq k \\ p_{k,\delta} (x) & k \leq x \leq k + \delta \\ (x - k) - \frac{1}{2}\delta & x \geq k + \delta \end{cases},$$
$$q_{k,\delta}^+ (x) := \int_k^x \eta_{k,\delta}^{+\ '} (\xi) f' (\xi) \, \mathrm{d}\xi.$$

Then $(\eta_{k,\delta}^+, q_{k,\delta}^+)$ is a boundary entropy pair and we have uniformly in x

$$|\eta_{k,\delta}^+ (x) - \eta_k^+ (x)| \leq \delta,$$
$$|\eta_{k,\delta}^{+\ '} (x) - \text{sign}^+ (x - k)| \to 0,$$
$$|q_{k,\delta}^+ (x) - q_k (x)| \leq \delta \, \|f'\|_{L^\infty([k,k+\delta])}.$$

Since $(\eta_{k,\delta}^+, q_{k,\delta}^+)$ is a boundary entropy pair we can put it into (6). For $\delta \to 0$ we obtain (7). In a similar way we can prove (7) for (η_k^-, q_k^-).

The following lemma concerns the trace of an entropy process solution and follows the ideas of Otto (see [14]).

Lemma 2.9 Let $\mu \in L^\infty(\Omega_T \times]0, 1[)$, $\mu = \mu(x, t, \alpha)$, be an entropy process solution of (1). Then for all $v \in L^\infty(]0, T[)$ and for all $\beta \in L^1(]0, T[)$, $\beta \geq 0$ a.e., we have:

$$\operatorname*{ess\,lim}_{x \uparrow x_r} \int_0^T \int_0^1 \operatorname{sign}^\pm(\mu - v)\,(f(\mu) - f(v))\,\beta \,\mathrm{d}\alpha \,\mathrm{dt}$$

$$\geq -\operatorname*{Lip}_{[-C,C]}(f) \int_0^T (u_r(t) - v)^\pm \beta \,\mathrm{dt},$$

$$\operatorname*{ess\,lim}_{x \downarrow x_l} \int_0^T \int_0^1 \operatorname{sign}^\pm(\mu - v)\,(f(\mu) - f(v))\,\beta \,\mathrm{d}\alpha \,\mathrm{dt}$$

$$\leq \operatorname*{Lip}_{[-C,C]}(f) \int_0^T (u_l(t) - v)^\pm \beta \,\mathrm{dt}.$$

The essential limits exist.

Proof Let $w \in \mathbb{Q}$ be fixed, (η_w^+, q_w^+) be a semi-Kružkov entropy pair and let $\beta \in C_0^\infty(]0, T[)$, $\beta \geq 0$, and $\psi \in C_0^\infty(\Omega)$, $\psi \geq 0$. Then we obtain with Lemma 2.8:

$$-\int_\Omega \int_0^T \int_0^1 q_w^+(\mu)\,\beta(t)\,\mathrm{d}\alpha\,\mathrm{dt}\psi'(x)\,\mathrm{dx}$$

$$\leq \int_0^T \int_\Omega \int_0^1 \eta_w^+(\mu)\,\beta'(t)\,\psi(x)$$

$$- \operatorname{sign}^+(\mu - w)\,z'(x)\,b(\mu)\,\beta(t)\,\psi(x)\,\mathrm{d}\alpha\,\mathrm{dx}\,\mathrm{dt} \tag{8}$$

$$\leq C \int_\Omega \psi(x)\,\mathrm{dx} \leq -C \int_\Omega x\psi'(x)\,\mathrm{dx}$$

where $C = C(\mu, z', b, \beta, w, T)$. But this implies the existence of a set E' of measure zero such that

$$x \mapsto \int_0^T \int_0^1 q_w^+(\mu)\,\beta(t)\,\mathrm{d}\alpha\,\mathrm{dt} - Cx \quad (x \in \Omega) \tag{9}$$

is non-increasing in $\Omega \setminus E'$. Furthermore there exist a set E'' of measure zero which only depends on μ and w, so that the function in (9) is bounded on $\Omega \setminus E''$. Therefore the function in (9) is monotone and bounded on $\Omega \setminus (E' \cup E'')$ and the essential limit exists, i.e.

$$\operatorname*{ess\,lim}_{x \uparrow x_r} \int_0^T \int_0^1 q_w^+(\mu)\,\beta(t)\,\mathrm{d}\alpha\,\mathrm{dt} = \lim_{\substack{x \uparrow x_r \\ x \notin E' \cup E''}} \int_0^T \int_0^1 q_w^+(\mu)\,\beta(t)\,\mathrm{d}\alpha\,\mathrm{dt}. \tag{10}$$

Similar as in (8) we get for $\psi \in C_0^\infty(]x_l, x_r[)$:

$$-\int_\Omega \int_0^T \int_0^1 q_w^+(\mu)\,\beta(t)\,\mathrm{d}\alpha\,\mathrm{dt}\psi'(x)\,\mathrm{dx}$$

$$\leq C \int_\Omega \psi(x)\,\mathrm{dx} + \operatorname*{Lip}_{[-C,C]}(f) \int_0^T \eta_w^+(u_r(t))\,\beta(t)\,\mathrm{dt}\psi(x_r). \tag{11}$$

By convolution with a nonnegative kernel we can also apply (11) to $\psi_\epsilon(x) := \frac{1}{\epsilon}(x - (x_r - \epsilon))^+$, $\epsilon > 0$. Therefore we obtain

$$
\begin{aligned}
&- \operatorname*{ess\,lim}_{x \uparrow x_r} \int_0^T \int_0^1 q_w^+(\mu) \, \beta(t) \, \mathrm{d}\alpha \, \mathrm{d}t \\
&= - \lim_{\epsilon \downarrow 0} \frac{1}{\epsilon} \int_{x_r - \epsilon}^{x_r} \int_0^T \int_0^1 q_w^+(\mu) \, \beta(t) \, \mathrm{d}\alpha \, \mathrm{d}t \, \mathrm{d}x \\
&\leq \operatorname*{Lip}_{[-C,C]}(f) \int_0^T \eta_w^+(u_r(t)) \, \beta(t) \, \mathrm{d}t.
\end{aligned}
\tag{12}
$$

By further approximation we can also get (12) for $\beta \in L^1(]0, T[)$, $\beta \geq 0$ a.e. Now let $v \in L^\infty(]0, T[)$, such that v has only a finite number of values $w_i \in \mathbb{Q}$, i.e.

$$
v = \sum_{i=1}^N w_i 1_{A_i}, \quad A_i \cap A_j = \emptyset \ (i \neq j), \quad \bigcup_{i=1,\dots,N} A_i =]0, 1[.
\tag{13}
$$

Let $\beta \in L^1(]0, T[)$ and $\beta_i := 1_{A_i}\beta$. Now using w_i and β_i in (12) and summing over i we obtain:

$$
\begin{aligned}
&\operatorname*{ess\,lim}_{x \uparrow x_r} \int_0^T \int_0^1 \operatorname{sign}^\pm(\mu - v(t)) \, (f(\mu) - f(v(t))) \, \beta(t) \, \mathrm{d}\alpha \, \mathrm{d}t \\
&\geq - \operatorname*{Lip}_{[-C,C]}(f) \int_0^T (u_r(t) - v(t))^+ \, \beta(t) \, \mathrm{d}t.
\end{aligned}
\tag{14}
$$

Since any L^1-function v is the limit of such elementary functions of type (13) we obtain (14) also for $v \in L^\infty(]0, T[)$.

For the semi-Kružkov entropy pairs (η_w^-, q_w^-) and the limits $x \downarrow x_l$ we proceed analogously.

Now we have to show that the entropy process solution has the correct initial data. Again the proof is similar to the one by Otto without source terms (see [12], Chapter 2, Lemma 7.41).

Lemma 2.10 Let $\mu \in L^\infty(\Omega_T \times]0, 1[)$, $\mu = \mu(x, t, \alpha)$, be an entropy process solution of (1). Then we have:

$$
\operatorname*{ess\,lim}_{t \downarrow 0} \int_\Omega \int_0^1 |\mu(x, t, \alpha) - u_0(x)| \, \mathrm{d}\alpha \, \mathrm{d}x = 0.
$$

Proof Let $\psi \in C_0^\infty(\Omega)$, choose (η_w^+, q_w^+) and (η_w^-, q_w^-) in (7) and for $\epsilon > 0$ use

$$
\varphi_\epsilon(x, t) := \left[1 - \frac{1}{\epsilon}t\right] 1_{[0,\epsilon]}(t) \, \psi(x)
$$

as test function. Adding those two inequalities we obtain

$$
\begin{aligned}
&\frac{1}{\epsilon} \int_0^\epsilon \int_\Omega \int_0^1 -|\mu - w| \, \psi(x) + F(\mu, w) \, (\epsilon - t) \, \psi'(x) \, \mathrm{d}\alpha \, \mathrm{d}x \, \mathrm{d}t \\
&\quad - \frac{1}{\epsilon} \int_0^\epsilon \int_\Omega \int_0^1 \operatorname{sign}(\mu - w) \, z'(x) \, b(\mu) \, (\epsilon - t) \, \psi(x) \, \mathrm{d}\alpha \, \mathrm{d}x \, \mathrm{d}t \\
&\quad + \int_\Omega |u_0(x) - w| \, \psi(x) \, \mathrm{d}x \geq 0.
\end{aligned}
$$

7

where

$$F(a, b) := \operatorname{sign}(a - b)(f(a) - f(b)) \quad \forall a, b \in \mathbb{R}. \tag{15}$$

Now we take the limes inferior with respect to ϵ and obtain

$$\operatorname{ess} \limsup_{t \downarrow 0} \int_\Omega \int_0^1 |\mu - w| \, \psi(x) \, \mathrm{d}\alpha \, \mathrm{dx} \leq \int_\Omega |u_0(x) - w| \, \psi(x) \, \mathrm{dx}. \tag{16}$$

Similar as in (14) we get that (16) for all $w \in \mathbb{R}$ implies

$$\operatorname{ess} \limsup_{t \downarrow 0} \int_\Omega \int_0^1 |\mu - v| \, \psi(x) \, \mathrm{d}\alpha \, \mathrm{dx} \leq \int_\Omega |u_0(x) - v| \, \psi(x) \, \mathrm{dx} \tag{17}$$

for all $v \in L^\infty(\Omega)$ and all $\psi \in L^1(\Omega)$, $\psi \geq 0$ a.e.

Choosing u_0 for v in (17) and using $\psi = 1$ as test function we get the statement of the lemma.

The proofs of the following two lemmata are similar to those in [15] and therefore we omit them. They are mainly based on Lemma 2.9.

Lemma 2.11 Let $\mu \in L^\infty(\Omega_T \times]0, 1[)$ be an entropy process solution of (1). Using F given by (15) we have for all $\varphi \in C_0^\infty(\overline{\Omega} \times]0, T[)$, $\varphi \geq 0$:

$$\int_0^T \int_\Omega \int_0^1 |\mu - k| \, \partial_t \varphi + F(\mu, k) \, \partial_x \varphi \, \mathrm{d}\alpha \, \mathrm{dx} \, \mathrm{dt}$$
$$- \int_0^T \int_\Omega \int_0^1 \operatorname{sign}(\mu - k) \, z'(x) \, b(\mu) \, \varphi \, \mathrm{d}\alpha \, \mathrm{dx} \, \mathrm{dt}$$
$$+ \operatorname{ess} \lim_{x \uparrow x_r} \int_0^T \int_0^1 F(\mu, u_r(t)) \, \varphi(x_r, t) \, \mathrm{d}\alpha \, \mathrm{dt}$$
$$- \operatorname{ess} \lim_{x \downarrow x_l} \int_0^T \int_0^1 F(\mu, u_l(t)) \, \varphi(x_l, t) \, \mathrm{d}\alpha \, \mathrm{dt}$$
$$- \int_0^T F(u_r(t), k) \, \varphi(x_r, t) \, dt + \int_0^T F(u_l(t), k) \, \varphi(x_l, t) \, dt \geq 0,$$

where $\mu = \mu(x, t, \alpha)$ and $\varphi = \varphi(x, t)$.

Lemma 2.12 Let $\mu, \nu \in L^\infty(\Omega_T \times]0, 1[)$ be two entropy process solutions of (1) with respect to the initial data $u_0 \in L^\infty(\Omega)$ and the boundary data $u_l, u_r \in L^\infty(]0, T[)$. Then donoting $\mu = \mu(x, t, \alpha)$, $\nu = \nu(x, t, \beta)$ we have for all nonnegative $\varphi \in C_0^\infty(\overline{\Omega} \times]0, T[)$, $\varphi = \varphi(x, t)$:

$$\int_{\Omega_T} \int_0^1 \int_0^1 |\mu - \nu| \, \partial_t \varphi + F(\mu, \nu) \, \partial_x \varphi \, \mathrm{d}\beta \, \mathrm{d}\alpha \, \mathrm{d}(x, t)$$
$$- \int_{\Omega_T} \int_0^1 \int_0^1 \operatorname{sign}(\mu - \nu) \, z'(x) \, (b(\mu) - b(\nu)) \, \varphi \, \mathrm{d}\beta \, \mathrm{d}\alpha \, \mathrm{d}(x, t) \geq 0, \tag{18}$$

where F is given by (15).

Now we are ready to prove the uniqueness of the entropy process solution. The proof without the source term can be found in [15], Theorem 2. For controlling the source term we use similar ideas as in [1], Theorem 5.

Proof 2.13 (Proof of Theorem 2.6) Let $\psi \in C_0^\infty(]0, T[)$, $\psi \geq 0$. In (18) we choose $(x, t) \mapsto \psi(t)$ as a test function and obtain with $\mu = \mu(x, t, \alpha)$, $\nu = \nu(x, t, \beta)$:

$$\int_{\Omega_T} \int_0^1 \int_0^1 |\mu - \nu| \, \partial_t \psi(t) \, \mathrm{d}\beta \, \mathrm{d}\alpha \, \mathrm{d}(x, t)$$
$$- \int_{\Omega_T} \int_0^1 \int_0^1 \mathrm{sign}\,(\mu - \nu) \, z'(x) \, (b(\mu) - b(\nu)) \, \psi(t) \, \mathrm{d}\beta \, \mathrm{d}\alpha \, \mathrm{d}(x, t) \geq 0. \tag{19}$$

Using

$$g(t) := \int_\Omega \int_0^1 \int_0^1 |\mu(x, t, \alpha) - \nu(x, t, \beta)| \, \mathrm{d}\beta \, \mathrm{d}\alpha \, \mathrm{d}x \quad \text{for } t \in]0, T[,$$

(19) can be written as

$$- \int_0^T g(t) \, \partial_t \psi(t) \, \mathrm{d}t \leq \|z'\|_{L^\infty(\Omega)} \|b'\|_{L^\infty(\mathbb{R})} \int_0^T g(t) \, \psi(t) \, \mathrm{d}t.$$

Integration by parts on the right hand side implies that there exists a set $E \subset]0, T[$ of measure zero, such that

$$t \mapsto g(t) - \|z'\|_{L^\infty(\Omega)} \|b'\|_{L^\infty(\mathbb{R})} \int_0^t g(\xi) \, \mathrm{d}\xi$$

is non-increasing on $]0, T[\backslash E$, i.e. for all $t_1, t_2 \in]0, T[\backslash E$, $t_1 < t_2$, we have:

$$g(t_2) \leq g(t_1) + \|z'\|_{L^\infty(\Omega)} \|b'\|_{L^\infty(\mathbb{R})} \int_{t_1}^{t_2} g(\xi) \, \mathrm{d}\xi. \tag{20}$$

Now we can apply the Gronwall Lemma and (20) implies

$$\|\mu(\cdot, t_2, \cdot) - \nu(\cdot, t_2, \cdot)\|_{L^1\left(\Omega \times]0, 1[^2\right)}$$
$$\leq \|\mu(\cdot, t_1, \cdot) - \nu(\cdot, t_1, \cdot)\|_{L^1\left(\Omega \times]0, 1[^2\right)} e^{\|z'\|_{L^\infty(\Omega)} \|b'\|_{L^\infty(\mathbb{R})} (t_2 - t_1)}. \tag{21}$$

Since the entropy process solution respects the initial data in the sense of Lemma 2.10 we have:

$$\mathrm{ess} \limsup_{t_1 \downarrow 0} \|\mu(\cdot, t_1, \cdot) - \nu(\cdot, t_1, \cdot)\|_{L^1\left(\Omega \times]0, 1[^2\right)}$$
$$\leq \mathrm{ess} \lim_{t_1 \downarrow 0} \left[\|\mu(\cdot, t_1, \cdot) - u_0\|_{L^1(\Omega \times]0, 1[)} + \|\nu(\cdot, t_1, \cdot) - u_0\|_{L^1(\Omega \times]0, 1[)} \right]$$
$$= 0,$$

and therefore from (21) for almost all $t_2 \in]0, T[$:

$$\|\mu(\cdot, t_2, \cdot) - \nu(\cdot, t_2, \cdot)\|_{L^1\left(\Omega \times]0, 1[^2\right)} \leq 0.$$

But this proves $\mu(x, t, \alpha) = \nu(x, t, \beta)$ for almost all $(x, t) \in \Omega_T$ and almost all $\alpha, \beta \in]0, 1[$. Therefore μ and ν do not depend on α and β respectively. Setting

$$u(x, t) := \int_0^1 \mu(x, t, \alpha) \, d\alpha$$

we get for almost all $(x, t, \alpha) \in \Omega_T \times]0, 1[$:

$$\mu(x, t, \alpha) = u(x, t) = \nu(x, t, \alpha).$$

Then, by Definitions 2.3 and 2.5, we see that u is an entropy solution of (1).

3   The Well-Balanced Scheme

In this section we will describe the well-balanced scheme for the initial boundary value problem (1). It was originally developed by Greenberg et al. in [9] for the initial value problem. For the Engquist-Osher numerical flux Perthame et al. (see [3]) could prove convergence of the numerical solution to the entropy solution. Here we are going to generalize this result to the initial boundary value problem (1).

First let us fix the notation for the discretization. Let $\Delta x > 0$ such that $\frac{|\Omega|}{\Delta x} \in \mathbb{N}$, $\Omega =]x_l, x_r[$, $j_l \in \mathbb{Z}$ and $x_j := x_l + (j - j_l + \frac{1}{2})\Delta x$, $C_j :=]x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}[$, choose $J \subset \mathbb{Z}$ such that $\overline{\Omega} = \bigcup_{j \in J} \overline{C}_j$ and define $j_r := \max J$.

Let $\Delta t > 0$ such that $N_T := \frac{T}{\Delta t} \in \mathbb{N}$ and $t^n := n\Delta t$. For any function $\varphi \in C^0(\mathbb{R}^2)$ we define $\varphi_j^n := \varphi(x_j, t^n)$ and the piecewise constant function

$$\overline{\varphi}(x, t) := \varphi_j^n \quad t \in [t^n, t^{n+1}), \ x \in C_j.$$

The main idea for the well-balanced schemes consists in the following fact. If v is a stationary solution of (1) then v satifies

$$\partial_x(D(v(x)) + z(x)) = 0 \quad \text{for a.a. } x \in \mathbb{R},$$

(for the definition of D see (3)) which is equivalent to

$$D(v) + z(x) = c \in \mathbb{R} \quad \text{for all } x \in \mathbb{R}. \tag{22}$$

This property is the main building block for the numerical scheme.

The standard form of a numerical scheme in conservation form for the partial differential equation in (1) is

$$u_j^{n+1} := u_j^n - \frac{\Delta t}{\Delta x} \left(g\left(u_j^n, u_{j+1}^n\right) - g\left(u_{j-1}^n, u_j^n\right) \right) - \Delta t z_j' u_j^n \quad \text{for } j \in J \qquad (23)$$

where g is a numerical flux and z_j' a discretisation of z' in C_j, e.g. the average of z' on the cell.

This scheme is very inefficient, especially for the approximation of stationary solutions. The well-balanced schemes are much better and are defined as follows.

Definition 3.1 (Well-Balanced Scheme) Let

$$u_j^0 := \frac{1}{\Delta x} \int_{C_j} u_0\left(x\right) \, \mathrm{dx} \qquad (24)$$

and define u_j^n for $j \notin J$ by

$$u_j^n = \begin{cases} u_l^n = \dfrac{1}{\Delta t} \displaystyle\int_{t^n}^{t^{n+1}} u_l\left(t\right) \, \mathrm{dt} & \text{for } j < j_l, \\[2mm] u_r^n = \dfrac{1}{\Delta t} \displaystyle\int_{t^n}^{t^{n+1}} u_r\left(t\right) \, \mathrm{dt} & \text{for } j > j_r. \end{cases} \qquad (25)$$

Similarly let z_j be given by

$$z_j := \begin{cases} z_{j_l} & \text{for } j < j_l, \\[2mm] \dfrac{1}{\Delta x} \displaystyle\int_{C_j} z\left(x\right) \, dx & \text{for } j \in J, \\[2mm] z_{j_r} & \text{for } j > j_r. \end{cases}$$

Now assume that $(u_j^n)_{j \in \mathbb{Z}}$ is already defined. Then the values u_j^{n+1} for the new time step of the well-balanced scheme are given by

$$u_j^{n+1} := u_j^n - \frac{\Delta t}{\Delta x} \left(g\left(u_j^n, u_{j+1,-}^n\right) - g\left(u_{j-1,+}^n, u_j^n\right) \right) \quad \text{for } j \in J \qquad (26)$$

where, due to (22), $u_{j+1,-}^n$ and $u_{j-1,+}^n$ are defined by

$$\begin{aligned} D\left(u_{j-1,+}\right) + z_j &= D\left(u_{j-1}\right) + z_{j-1}, \\ D\left(u_{j+1,-}\right) + z_j &= D\left(u_{j+1}\right) + z_{j+1}. \end{aligned} \qquad (27)$$

Using the discrete data $(u_j^n)_{j \in \mathbb{Z}, 0 \leq n < N_T}$ we define the numerical solution $u_{\Delta x} \in L^\infty(\mathbb{R} \times [0, T[)$ by

$$u_{\Delta x}\left(x, t\right) := u_j^n \quad \text{for } t \in [t^n, t^{n+1}[, \ n < N_T, \ x \in C_j, \ j \in \mathbb{Z}. \qquad (28)$$

Remark 3.2 Under condition (4) there always exist unique solutions $u_{j-1,+}$ and $u_{j+1,-}$ of (27).

Remark 3.3 If the values u_j^n are "stationary", i.e. $D(u_j^n) + z_j = D(u_{j-1}^n) + z_{j-1}$ for all $j \in J$, then we obtain

$$u_{j-1,+}^n = u_j^n, \quad u_{j+1,-}^n = u_j^n$$

as solutions of (27) and therefore $u_j^{n+1} = u_j^n$ for all $j \in J$.

Now we assume that the numerical flux is given by the Engquist-Osher flux

$$g(u,v) = \int_0^u f'^+(\xi) \, d\xi - \int_0^v f'^-(\xi) \, d\xi + f(0) \quad \text{for } u, v \in \mathbb{R}. \tag{29}$$

and set

$$H(u,v,w) := v - \frac{\Delta t}{\Delta x}(g(v,w) - g(u,v)) \quad \text{for } u, v, w \in \mathbb{R}. \tag{30}$$

Before proving the boundedness of the numerical solutions we will prove that the numerical values obtained by the Engquist-Osher scheme remain in the convex hull of its data.

As in [3] we use the following "density-function" in order to move the integration bounds to the integrand:

$$\chi_s(\xi) := \begin{cases} 1 & \text{for } 0 < \xi < s, \\ -1 & \text{for } s < \xi < 0, \\ 0 & \text{for } otherwise. \end{cases}$$

Now we can write the Engquist-Osher scheme (30) as an integral by defining

$$\begin{aligned} h(u,v,w)(\xi) := \chi_v(\xi) &- \frac{\Delta t}{\Delta x}\left(f'^+(\xi)\chi_v(\xi) - f'^-(\xi)\chi_w(\xi)\right) \\ &+ \frac{\Delta t}{\Delta x}\left(f'^+(\xi)\chi_u(\xi) + f'^-(\xi)\chi_v(\xi)\right) \end{aligned} \tag{31}$$

for all $u, v, w, \xi \in \mathbb{R}$. Then (30) can be written as

$$H(u,v,w) = \int_{\mathbb{R}} h(u,v,w)(\xi) \, d\xi \quad \text{for } u, v, w \in \mathbb{R}.$$

This form can be used to show that the values produced by the Engquist-Osher scheme are the convex hull of its arguments.

Lemma 3.4 Let $h : \mathbb{R}^3 \to L^1(\mathbb{R})$ be defined by (31), $K = [-C, C]$, $u, v, w \in K$ and assume that the CFL-condition

$$\|f'\|_{L^\infty(K)} \frac{\Delta t}{\Delta x} \le 1$$

holds. Then for all $\xi \in \mathbb{R}$ we have

$$h(u,v,w)(\xi) \in \text{conv}\{\chi_u(\xi), \chi_v(\xi), \chi_w(\xi)\},$$
$$0 \le \text{sign}(\xi) h(u,v,w)(\xi) \le 1$$

and consequently

$$|H(u,v,w)| \le \max\{|u|, |v|, |w|\}. \tag{32}$$

Proof The proof uses only standard means, see also [3].

In the next step we will prove the boundedness of the numerical solutions.

Lemma 3.5 (Boundedness of the Numerical Solutions) Let (2), (4), (29) be satisfied,

$$M := \max \left\{ \|u_0\|_{L^\infty(\Omega)}, \|u_l\|_{L^\infty(]0,T[)}, \|u_r\|_{L^\infty(]0,T[)} \right\}, \tag{33}$$

$$\begin{aligned}
C_T^{\Delta x} &:= M e^{2T\|b'\|_{L^\infty(\mathbb{R})}\|z'\|_{L^\infty(\Omega)}} \\
&\quad + \Delta x \frac{\|z'\|_{L^\infty(\Omega)}}{\inf_\mathbb{R} D'} e^{4T\|b'\|_{L^\infty(\mathbb{R})}\|z'\|_{L^\infty(\Omega)}} \\
&\quad + |b(0)| e^{2T(\|b'\|_{L^\infty(\mathbb{R})}+1)\|z'\|_{L^\infty(\Omega)}},
\end{aligned} \tag{34}$$

$$K_T^{\Delta x} := \left[-C_T^{\Delta x}, C_T^{\Delta x} \right]$$

and we assume the CFL-condition $\operatorname{Lip}_{K_T^{\Delta x}}(f) \frac{\Delta t}{\Delta x} \leq 1$. *Then we obtain for all* $n\,\Delta t \leq T$:

$$\sup_{j \in J} |u_j^n| \leq C_T^{\Delta x}. \tag{35}$$

Furthermore we have for all $(n+1)\,\Delta t \leq T$:

$$\sup_{j \in J} |u_{j-1,+}^n| \leq C_T^{\Delta x}, \quad \sup_{j \in J} |u_{j+1,-}^n| \leq C_T^{\Delta x}. \tag{36}$$

Proof Let $u, v, w \in \mathbb{R}$ be arbitrary. Then we have with $b_0 := b(0)$:

$$\begin{aligned}
|g(u,v) - g(u,w)| = \left| \int_v^w f'^-(s)\,\mathrm{d}s \right| &\leq \int_{[v,w]} \left| f'^-(s) \right| \mathrm{d}s \\
&\leq \int_{[v,w]} |f'(s)|\,\mathrm{d}s = \int_{[v,w]} |D'(s)\,b(s)|\,\mathrm{d}s \\
&\leq \|b'\|_\infty \int_{[v,w]} D'(s)\,|s|\,\mathrm{d}s + |b_0| \int_{[v,w]} D'(s)\,\mathrm{d}s \\
&\leq [\|b'\|_\infty \max\{|v|,|w|\} + |b_0|] \int_{[v,w]} D'(s)\,\mathrm{d}s \\
&\leq [\|b'\|_\infty \max\{|v|,|w|\} + |b_0|]\,|D(v) - D(w)|.
\end{aligned}$$

The mean value theorem implies:

$$\begin{aligned}
\left| u_{j+1}^n - u_{j+1,-}^n \right| &\leq \frac{|D(u_{j+1}^n) - D(u_{j+1,-}^n)|}{\inf_\mathbb{R} D'} \\
&= \frac{|z_j - z_{j+1}|}{\inf_\mathbb{R} D'} \leq \Delta x \frac{\|z'\|_\infty}{\inf_\mathbb{R} D'}
\end{aligned} \tag{37}$$

and therefore:

$$\begin{aligned}
&\left| g\left(u_j^n, u_{j+1,-}^n\right) - g\left(u_j^n, u_{j+1}^n\right) \right| \\
&\leq \left[\|b'\|_\infty \left(|u_{j+1}^n| + \Delta x \frac{\|z'\|_\infty}{\inf_\mathbb{R} D'} \right) + |b_0| \right] \Delta x\, \|z'\|_\infty.
\end{aligned}$$

13

In a similar way we obtain:

$$\left| g\left(u_{j-1,+}^n, u_j^n\right) - g\left(u_{j-1}^n, u_j^n\right)\right|$$
$$\leq \left[\|b'\|_\infty \left(\left|u_{j-1}^n\right| + \Delta x \frac{\|z'\|_\infty}{\inf_{\mathbb{R}} D'}\right) + |b_0|\right] \Delta x \|z'\|_\infty.$$

Now by induction with respect to $n \in \mathbb{N}$ such that $n\Delta t \leq T$ we will show:

$$\sup_{j \in J} \left|u_j^n\right| \leq M e^{2n\Delta t \|b'\|_\infty \|z'\|_\infty} + \Delta x \frac{\|z'\|_\infty}{\inf_{\mathbb{R}} D'} e^{4n\Delta t \|b'\|_\infty \|z'\|_\infty}$$
$$+ |b_0| \, e^{2n\Delta t \left(\|b'\|_\infty + 1\right)\|z'\|_\infty}, \tag{38}$$

which will prove the statement (35) of the lemma. Obviously (38) holds for the initial data, i.e. $n = 0$.

Assume that (38) holds for $n \in \mathbb{N}$, $(n+1)\Delta t \leq T$. Since the maximum principle holds for the Engquist-Osher flux $g(u, v)$ (see (32)), the CFL-condition holds and $\mathrm{Lip}_{K_T^{\Delta x}}(g) \leq \mathrm{Lip}_{K_T^{\Delta x}}(f)$ we obtain for all $j \in J$:

$$\left|u_j^{n+1}\right| \leq \left|u_j^n - \frac{\Delta t}{\Delta x}\left(g\left(u_j^n, u_{j+1}^n\right) - g\left(u_{j-1}^n, u_j^n\right)\right)\right|$$
$$+ \left|\frac{\Delta t}{\Delta x}\left(g\left(u_j^n, u_{j+1,-}^n\right) - g\left(u_j^n, u_{j+1}^n\right)\right)\right|$$
$$+ \left|\frac{\Delta t}{\Delta x}\left(g\left(u_{j-1,+}^n, u_j^n\right) - g\left(u_{j-1}^n, u_j^n\right)\right)\right|$$
$$\leq \sup_{j \in \mathbb{Z}} \left|u_j^n\right| + 2\Delta t \left[\|b'\|_\infty \left[\sup_{j \in \mathbb{Z}} \left|u_j^n\right| + \frac{\Delta x \|z'\|_\infty}{\inf_{\mathbb{R}} D'}\right] + |b_0|\right] \|z'\|_\infty$$
$$\leq \max\left\{\sup_{j \in J} \left|u_j^n\right|, |u_l^n|, |u_r^n|\right\} \left(1 + 2\Delta t \|b'\|_\infty \|z'\|_\infty\right)$$
$$+ 2\Delta t |b_0| \|z'\|_\infty + 2\Delta x \Delta t \|b'\|_\infty \frac{\|z'\|_\infty^2}{\inf_{\mathbb{R}} D'}.$$

If $\max\{\sup_{j \in J} |u_j^n|, |u_l^n|, |u_r^n|\} = |u_l^n|$, we have

$$\left|u_j^{n+1}\right| \leq \|u_l\|_{L^\infty(]0,T[)} \, e^{2\Delta t \|b'\|_\infty \|z'\|_\infty} \underbrace{e^{2n\Delta t \|b'\|_\infty \|z'\|_\infty}}_{\geq 1}$$
$$+ \frac{\Delta x \|z'\|_\infty}{\inf_{\mathbb{R}} D'} e^{4\Delta t \|b'\|_\infty \|z'\|_\infty} \underbrace{e^{4n\Delta t \|b'\|_\infty \|z'\|_\infty}}_{\geq 1}$$
$$+ |b_0| \, e^{2\Delta t \|z'\|_\infty} e^{2n\Delta t \left(\|b'\|_\infty + 1\right)\|z'\|_\infty}$$
$$\leq M e^{2(n+1)\Delta t \|b'\|_\infty \|z'\|_\infty} + \frac{\Delta x \|z'\|_\infty}{\inf_{\mathbb{R}} D'} e^{4(n+1)\Delta t \|b'\|_\infty \|z'\|_\infty}$$
$$+ |b_0| \, e^{2(n+1)\Delta t \left(\|b'\|_\infty + 1\right)\|z'\|_\infty}.$$

14

In a similar way the case $\max\{\sup_{j\in J}|u_j^n|,|u_l^n|,|u_r^n|\} = |u_r^n|$ can be handled.

If $\max\{\sup_{j\in J}|u_j^n|,|u_l^n|,|u_r^n|\} = \sup_{j\in J}|u_j^n|$ we use (38) to obtain:

$$
\begin{aligned}
\left|u_j^{n+1}\right| &\leq M e^{2n\Delta t\|b'\|_\infty\|z'\|_\infty}\left(1 + 2\Delta t\,\|b'\|_\infty\,\|z'\|_\infty\right) \\
&\quad + \frac{\Delta x\,\|z'\|_\infty}{\inf_{\mathbb{R}} D'}e^{4n\Delta t\|b'\|_\infty\|z'\|_\infty}\left(1 + 2\Delta t\,\|b'\|_\infty\,\|z'\|_\infty\right) \\
&\quad + |b_0|\,e^{2n\Delta t\left(\|b'\|_\infty+1\right)\|z'\|_\infty}\left(1 + 2\Delta t\,\|b'\|_\infty\,\|z'\|_\infty\right) \\
&\quad + 2\Delta t\,|b_0|\,\|z'\|_\infty + 2\Delta x\Delta t\,\|b'\|_\infty\,\frac{\|z'\|_\infty^2}{\inf_{\mathbb{R}} D'}. \\
&\leq M e^{2(n+1)\Delta t\|b'\|_\infty\|z'\|_\infty} + \frac{\Delta x\,\|z'\|_\infty}{\inf_{\mathbb{R}} D'}e^{4(n+1)\Delta t\|b'\|_\infty\|z'\|_\infty} \\
&\quad + |b_0|\,e^{2(n+1)\Delta t\left(\|b'\|_\infty+1\right)\|z'\|_\infty}.
\end{aligned}
$$

Therefore we have proved (38) for $n\Delta t \leq T$. The estimates in (36) now follow from (37), (38) and $(n+1)\,\Delta t \leq T$.

Notice that the second term in (34) is necessary, although it is missing in the proof for the initial value problem in [3], Lemma 3.1. This is shown in the diplom thesis [13].

A main tool for the convergence proof in Section 4 will be a cell entropy inequality. This means for an arbitrary entropy pair (η, q) we have to find a numerical entropy flux G, such that for $C > 0$ and for $\frac{\Delta t}{\Delta x}$ sufficiently small

$$\eta\left(H\left(u,v,w\right)\right) - \eta\left(v\right) + \frac{\Delta t}{\Delta x}\left(G\left(v,w\right) - G\left(u,v\right)\right) \leq 0 \tag{39}$$

holds for all $u, v, w \in [-C, C]$.

To prove the inequality in (39) we will need a lemma by Brenier (see [5]).

Lemma 3.6 (Brenier [5]) Let $f \in L^1\left(\mathbb{R}\right)$ such that $0 \leq \mathrm{sign}\left(\xi\right)f\left(\xi\right) \leq 1$ for all $\xi \in \mathbb{R}$ and let $h \in C^2\left(\mathbb{R}\right)$ be convex and Lipschitz-continuous. Then we have:

$$h\left(\int_{\mathbb{R}} f\left(\xi\right)\,\mathrm{d}\xi\right) - h\left(0\right) \leq \int_{\mathbb{R}} h'\left(\xi\right)f\left(\xi\right)\,\mathrm{d}\xi.$$

Now we can prove the cell entropy inequality (39), which generalizes a result due to Perthame et al., see [3].

Lemma 3.7 Let $|u|,|v|,|w| \leq C \in \mathbb{R}$, $K := [-C, C]$, (η, q) be an entropy pair and assume the CFL-condition

$$\|f'\|_{L^\infty(K)}\frac{\Delta t}{\Delta x} \leq 1. \tag{40}$$

Then $G : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ given by

$$G(u, v) := \int_0^u \eta'(\xi) f'^+(\xi) \, \mathrm{d}\xi - \int_0^v \eta'(\xi) f'^-(\xi) \, \mathrm{d}\xi + q(0) \tag{41}$$

is a consistent numerical entropy flux, such that the cell entropy inequality (39) is satisfied.

Proof It is easy to check that G is a consistent numerical entropy flux. Since $|u|, |v|, |w| \leq C$ and due to (40) $|H(u, v, w)| \leq C$, we can assume without loss of generality that η is Lipschitz-continuous.

Due to the CFL-condition (40) and Lemma 3.4 we have for all $\xi \in \mathbb{R}$ that

$$0 \leq \operatorname{sign}(\xi) h(u, v, w)(\xi) \leq 1$$

and by Lemma 3.6 we have

$$\eta \left(\int_{\mathbb{R}} h(u, v, w)(\xi) \, \mathrm{d}\xi \right) - \eta(0) \leq \int_{\mathbb{R}} \eta'(\xi) h(u, v, w)(\xi) \, \mathrm{d}\xi.$$

Then using (31) and (41) on the right hand side, the cell entropy inequality (39) follows.

4 Main Result and Convergence of the Numerical Solutions

In this section we will present the main result of this paper, the convergence of the numerical solutions of the well-balanced scheme (see Definition 3.1) to the entropy solution, and its proof.

Theorem 4.1 Let (2), (4) and (29) be satisfied, let $(u_{\Delta x})_{\Delta x > 0}$ be the sequence of numerical solutions of (1), in the sense of Definition 3.1 and assume the CFL-condition $\operatorname{Lip}_{K_T^{\Delta x}}(f) \frac{\Delta t}{\Delta x} \leq 1$, where $C_T^{\Delta x}$ is defined as in (34). Then there exists a $u \in L^\infty(\Omega_T)$, such that

$$u_{\Delta x} \xrightarrow{\Delta x \to 0} u \quad in \ L^p(\Omega_T) \quad for \ all \ 1 \leq p < \infty \tag{42}$$

and u is the unique entropy solution of (1).

Proof see Proof 4.5

For the proof we need the following lemmata.

Lemma 4.2 Let (η, q) be a boundary entropy pair, let $w \in \mathbb{R}$ such that $\eta(w) = \eta'(w) = q(w) = 0$ and let G be defined by (41). Then we have for all $u, v \in \mathbb{R}$

$$G(u, v) = \int_w^u \eta'(\xi) f'^+(\xi) \, \mathrm{d}\xi - \int_w^v \eta'(\xi) f'^-(\xi) \, \mathrm{d}\xi. \tag{43}$$

If additionally $u, v, w \in [-C, C]$, $C > 0$, holds, the following inequality holds true:

$$- \operatorname*{Lip}_{[-C,C]}(f) \eta(v) \leq G(u, v) \leq \operatorname*{Lip}_{[-C,C]}(f) \eta(u). \tag{44}$$

16

Proof Since $q(w) = 0$ we have

$$q(0) = \int_w^0 q'(\xi)\,\mathrm{d}\xi = \int_w^0 \eta'(\xi)\,f'(\xi)\,\mathrm{d}\xi$$
$$= \int_w^0 \eta'(\xi)\,f'^+(\xi)\,\mathrm{d}\xi - \int_w^0 \eta'(\xi)\,f'^-(\xi)\,\mathrm{d}\xi.$$

Taking (41) into account, we obtain (43).

Since $\eta \in C^2(\mathbb{R})$ is convex and $\eta'(w) = 0$, we have

$$\mathrm{sign}\,(u - w)\,\eta'(\xi) \geq 0 \quad \text{for all } \xi \in [u, w].$$

Therefore, as $f'^+ \geq 0$, we see that

$$\int_w^u \eta'(\xi)\,f'^+(\xi)\,\mathrm{d}\xi = \int_{[w,u]} \mathrm{sign}\,(u - w)\,\eta'(\xi)\,f'^+(\xi)\,\mathrm{d}\xi \geq 0.$$

On the other hand $\eta(w) = 0$ and we can estimate:

$$\int_w^u \eta'(\xi)\,f'^+(\xi)\,\mathrm{d}\xi \leq \|f'\|_{L^\infty([-C,C])} \int_{[w,u]} \mathrm{sign}\,(u - w)\,\eta'(\xi)\,\mathrm{d}\xi$$
$$= \mathrm{Lip}_{[-C,C]}\,(f)\,\eta(u).$$

Applying the same arguments we see that

$$0 \leq \int_w^v \eta'(\xi)\,f'^-(\xi)\,\mathrm{d}\xi \leq \mathrm{Lip}_{[-C,C]}\,(f)\,\eta(v)$$

and using (43) we obtain (44).

Lemma 4.3 Let (2), (4), (29) and the cell entropy inequality (39) be satisfied. For $\Delta t, \Delta x > 0$ assume that $\frac{\Delta t}{\Delta x} \leq \lambda$ and let $(u_j^n)_{0 \leq n \leq N_T, j \in J}$ be the numerical solution of (1) in the sense of Definition 3.1. Let (η, q) be a boundary entropy pair, $w \in \mathbb{R}$ such that $\eta(w) = \eta'(w) = q(w) = 0$, $|u_j^n|, |w| \leq C$ for all $j \in J$, $n \leq N_T$ and $|u_{j+1,-}^n|, |u_{j-1,+}^n| \leq C$ for all $j \in J$, $n < N_T$. Then for all $\varphi \in C_0^2(\mathbb{R}^2) \cap C_0^\infty(\overline{\Omega} \times [0, T))$, $\varphi \geq 0$, we have

$$\int_0^T \int_\Omega \eta(u_{\Delta x})\,\partial_t\varphi + q(u_{\Delta x})\,\partial_x\varphi - \eta'(u_{\Delta x})\,b(u_{\Delta x})\,z'(x)\,\varphi\,\mathrm{d}x\,\mathrm{d}t$$
$$+ \int_\Omega \eta(u_0(x))\,\varphi(x, 0)\,\mathrm{d}x + \mathrm{Lip}_{[-C,C]}\,(f) \int_0^T \eta(u_r(t))\,\varphi(x_{j_r}, t)\,\mathrm{d}t \qquad (45)$$
$$+ \mathrm{Lip}_{[-C,C]}\,(f) \int_0^T \eta(u_l(t))\,\varphi(x_{j_l}, t)\,\mathrm{d}t \geq \mathcal{O}(\Delta x).$$

Proof Multiplying the cell entropy inequality (39) by $\Delta x\, \varphi_j^n$ and summing over $n < N_T$ and $j \in J$ we obtain

$$
\begin{aligned}
0 \geq\ &\Delta x \sum_{n=0}^{N_T-1} \sum_{j \in J} \left(\eta\left(u_j^{n+1}\right) - \eta\left(u_j^n\right) \right) \varphi_j^n \\
&+ \Delta t \sum_{n=0}^{N_T-1} \sum_{j \in J} \left(G\left(u_j^n, u_{j+1}^n\right) - G\left(u_{j-1}^n, u_j^n\right) \right) \varphi_j^n \\
&+ \Delta t \sum_{n=0}^{N_T-1} \sum_{j \in J} \left(G\left(u_j^n, u_{j+1,-}^n\right) - G\left(u_j^n, u_{j+1}^n\right) \right) \varphi_j^n \\
&- \Delta t \sum_{n=0}^{N_T-1} \sum_{j \in J} \left(G\left(u_{j-1,+}^n, u_j^n\right) - G\left(u_{j-1}^n, u_j^n\right) \right) \varphi_j^n.
\end{aligned}
\tag{46}
$$

We will now treat the four sums in (46) seperately. The first term is handled in a standard way by an index shifting argument:

$$
\begin{aligned}
\sum_{n=0}^{N_T-1} &\left(\eta\left(u_j^{n+1}\right) - \eta\left(u_j^n\right) \right) \varphi_j^n \\
&= - \sum_{n=0}^{N_T-1} \eta\left(u_j^n\right) \left(\varphi_j^n - \varphi_j^{n-1} \right) + \eta\left(u_j^{N_T}\right) \varphi_j^{N_T-1} - \eta\left(u_j^0\right) \varphi_j^{-1} \\
&= -\Delta t \sum_{n=0}^{N_T-1} \eta\left(u_j^n\right) \left(\partial_t \varphi\right)_j^n - \eta\left(u_j^0\right) \varphi_j^0 + \mathcal{E}_{t,j},
\end{aligned}
\tag{47}
$$

where $|\mathcal{E}_{t,j}| \leq \Delta t\, \|\eta\|_{L^\infty([-C,C])} \left(T\, \|\partial_t^2 \varphi\|_\infty + 2\, \|\partial_t \varphi\|_\infty \right).$

Using Jensen's inequality we deduce that

$$
\eta\left(u_j^0\right) \leq \frac{1}{\Delta x} \int_{C_j} \eta\left(u_0\left(x\right)\right)\, \mathrm{dx}
$$

and from (47) we obtain

$$
\begin{aligned}
\sum_{n=0}^{N_T-1} &\left(\eta\left(u_j^{n+1}\right) - \eta\left(u_j^n\right) \right) \varphi_j^n \\
&\geq -\Delta t \sum_{n=0}^{N_T-1} \eta\left(u_j^n\right) \left(\partial_t \varphi\right)_j^n - \frac{1}{\Delta x} \int_{C_j} \eta\left(u_0\left(x\right)\right)\, \mathrm{dx} \varphi_j^0 + \mathcal{E}_{t,j}.
\end{aligned}
\tag{48}
$$

Now we turn to the second term of (46). Using (43), an index shifting argument and (44) we obtain

$$
\sum_{j \in J} \left(G\left(u_j^n, u_{j+1}^n\right) - G\left(u_{j-1}^n, u_j^n\right) \right) \varphi_j^n
$$

18

$$= -\Delta x \sum_{j \in J} \int_w^{u_j^n} \eta'(\xi) f'(\xi) \, d\xi \, (\partial_x \varphi)_j^n$$

$$+ G\left(u_{j_r}^n, u_{j_r+1}^n\right) \varphi_{j_r}^n - G\left(u_{j_l-1}^n, u_{j_l}^n\right) \varphi_{j_l}^n + \mathcal{E}_{x,n}$$

$$\geq -\Delta x \sum_{j \in J} q\left(u_j^n\right) (\partial_x \varphi)_j^n - \eta\left(u_{j_r+1}^n\right) \operatorname*{Lip}_{[-C,C]}(f) \varphi_{j_r}^n$$

$$- \eta\left(u_{j_l-1}^n\right) \operatorname*{Lip}_{[-C,C]}(f) \varphi_{j_l}^n + \mathcal{E}_{x,n}.$$

where $|\mathcal{E}_{x,n}| \leq c_{x,n} \Delta x$ with

$$c_{x,n} = 2C \, \|\eta'\|_{L^\infty([-C,C])} \, \|f'\|_{L^\infty([-C,C])} \left(|\Omega| \, \left\|\partial_x^2 \varphi\right\|_\infty + \|\partial_x \varphi\|_\infty\right).$$

Again using Jensen's inequality to get

$$\eta(u_l^n) \leq \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \eta(u_l(t)) \, dt, \quad \eta(u_r^n) \leq \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \eta(u_r(t)) \, dt,$$

we arrive at the following inequality:

$$\sum_{j \in J} \left(G\left(u_j^n, u_{j+1}^n\right) - G\left(u_{j-1}^n, u_j^n\right)\right) \varphi_j^n$$

$$\geq -\Delta x \sum_{j \in J} q\left(u_j^n\right) (\partial_x \varphi)_j^n - \operatorname*{Lip}_{[-C,C]}(f) \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \eta(u_r(t)) \varphi_{j_r}^n \, dt \qquad (49)$$

$$- \operatorname*{Lip}_{[-C,C]}(f) \frac{1}{\Delta t} \int_{t^n}^{t^{n+1}} \eta(u_l(t)) \varphi_{j_l}^n \, dt + \mathcal{E}_{x,n}.$$

Now we turn our attention to the discretisation of the source term. By the mean value theorem there is a $\zeta_{j+1}^n \in [u_{j+1}^n, u_{j+1,-}^n]$ such that

$$G\left(u_j^n, u_{j+1,-}^n\right) - G\left(u_j^n, u_{j+1}^n\right)$$

$$= -\int_{u_{j+1}^n}^{u_{j+1,-}^n} \eta'(\xi) f'^{-}(\xi) \, d\xi$$

$$= -\eta'\left(\zeta_{j+1}^n\right) f'^{-}\left(\zeta_{j+1}^n\right) \left(u_{j+1,-}^n - u_{j+1}^n\right).$$

Also by the mean value theorem and (27) there is a $\vartheta_{j+1}^n \in [u_{j+1}^n, u_{j+1,-}^n]$ so that

$$z_{j+1} - z_j = D\left(u_{j+1,-}^n\right) - D\left(u_{j+1}^n\right) = D'\left(\vartheta_{j+1}^n\right) \left(u_{j+1,-}^n - u_{j+1}^n\right)$$

and since $\inf_{\mathbb{R}} D' > 0$ we have

$$G\left(u_j^n, u_{j+1,-}^n\right) - G\left(u_j^n, u_{j+1}^n\right) = -\Delta x \eta'\left(\zeta_{j+1}^n\right) \frac{f'^{-}\left(\zeta_{j+1}^n\right)}{D'\left(\vartheta_{j+1}^n\right)} \frac{z_{j+1} - z_j}{\Delta x}. \qquad (50)$$

19

Since $z_{j_l} - z_{j_l-1} = z_{j_r+1} - z_{j_r} = 0$, multiplying (50) by φ_j^n and summing over $j \in J$ we obtain

$$\sum_{j \in J} \left(G\left(u_j^n, u_{j+1,-}^n\right) - G\left(u_j^n, u_{j+1}^n\right) \right) \varphi_j^n$$

$$= -\Delta x \sum_{j \in J} \eta'\left(\zeta_j^n\right) \frac{f'^-(\zeta_j^n)}{D'(\vartheta_j^n)} \frac{z_j - z_{j-1}}{\Delta x} \varphi_{j-1}^n \tag{51}$$

$$= -\Delta x \sum_{j \in J} \eta'\left(u_j^n\right) \frac{f'^-(u_j^n)}{D'(u_j^n)} \frac{z_j - z_{j-1}}{\Delta x} \varphi_j^n + \mathcal{E}_{+,n},$$

where

$$\mathcal{E}_{+,n} := \sum_{j \in J} \left[\eta'\left(u_j^n\right) \frac{f'^-(u_j^n)}{D'(u_j^n)} \varphi_j^n - \eta'\left(\zeta_j^n\right) \frac{f'^-(\zeta_j^n)}{D'(\vartheta_j^n)} \varphi_{j-1}^n \right] (z_j - z_{j-1}).$$

By the assumptions (4) $\frac{1}{D'}$ is differentiable and we can estimate:

$$\left| \eta'\left(\zeta_j^n\right) \frac{f'^-(\zeta_j^n)}{D'(\vartheta_j^n)} - \eta'\left(u_j^n\right) \frac{f'^-(u_j^n)}{D'(u_j^n)} \right|$$

$$\leq C^* \max\left\{ \left| \zeta_j^n - u_j^n \right|, \left| \vartheta_j^n - u_j^n \right| \right\} \leq C^* \Delta x \frac{\|z'\|_\infty}{\inf_{\mathbb{R}} D'}$$

where $C^* = C^*(C, \eta, f, D)$. Using this result we see that $|\mathcal{E}_{+,n}| \leq c_{\pm,n} \Delta x$ with

$$c_{\pm,n} = |\Omega| \, C^* \frac{\|z'\|_\infty^2}{\inf_{\mathbb{R}} D'} \|\varphi\|_\infty$$

$$+ |\Omega| \frac{\|\eta'\|_{L^\infty([-C,C])} \|f'\|_{L^\infty([-C,C])}}{\inf_{\mathbb{R}} D'} \|z'\|_\infty \|\varphi'\|_\infty.$$

In the same way we treat the other part of the source term and obtain

$$-\sum_{j \in J} \left(G\left(u_{j-1,+}^n, u_j^n\right) - G\left(u_{j-1}^n, u_j^n\right) \right) \varphi_j^n$$

$$= \Delta x \sum_{j \in J} \eta'\left(u_j^n\right) \frac{f'^+(u_j^n)}{D'(u_j^n)} \frac{z_{j+1} - z_j}{\Delta x} \varphi_j^n + \mathcal{E}_{-,n}, \tag{52}$$

where $|\mathcal{E}_{-,n}| \leq c_{\pm,n} \Delta x$.

Using (48), (49), (51) and (52) in (46) we arrive at the following inequality:

$$
\begin{aligned}
&\Delta t \Delta x \sum_{n=0}^{N_T-1} \sum_{j \in J} \eta\left(u_j^n\right)\left(\partial_t \varphi\right)_j^n + \Delta t \Delta x \sum_{n=0}^{N_T-1} \sum_{j \in J} q\left(u_j^n\right)\left(\partial_x \varphi\right)_j^n \\
&-\Delta t \Delta x \sum_{\substack{0 \le n < N_T \\ j \in J}} \eta'\left(u_j^n\right)\left[\frac{f'^{+}(u_j^n)}{D'(u_j^n)} \frac{z_{j+1}-z_j}{\Delta x} - \frac{f'^{-}(u_j^n)}{D'(u_j^n)} \frac{z_j - z_{j-1}}{\Delta x}\right] \varphi_j^n \\
&\qquad\qquad + \sum_{j \in J} \int_{C_j} \eta\left(u_0\left(x\right)\right) \varphi_j^0 \, \mathrm{dx} \\
&+ \operatorname*{Lip}_{[-C,C]}(f) \sum_{n=0}^{N_T-1} \int_{t^n}^{t^{n+1}} \eta\left(u_r\left(t\right)\right) \varphi_{jr}^n + \eta\left(u_l\left(t\right)\right) \varphi_{jl}^n \, \mathrm{dt} \ge \mathcal{O}\left(\Delta x\right).
\end{aligned}
\tag{53}
$$

It is not hard to see that (53) implies (45).

For the final convergence proof we need the following relation between weak$*$ and strong convergence:

Lemma 4.4 *Let Ω be a bounded, measurable subset of $\mathbb{R}^n$, let $(v_n)_{n \in \mathbb{N}}$ be a bounded sequence in $L^\infty(\Omega)$ and $v \in L^\infty(\Omega)$ such that for all continuous functions $g \in C(\mathbb{R})$ we have*

$$
g\left(v_n\right) \to g\left(v\right) \quad weak * \;\; in \;\; L^\infty\left(\Omega\right).
$$

Then the convergence $v_n \to v$ is strong in $L^p(\Omega)$ for $1 \le p < \infty$.

Proof see [15], Lemma 10.

Proof 4.5 (Proof of Theorem 4.1) Let $\delta > 0$ be fixed. Then for $\Delta x \le \delta$ we have $C_T^{\Delta x} \le C_T^\delta$, where $C_T^{\Delta x}$ is defined by (34). Hence the CFL-condition and Lemma 3.5 provide for $\Delta x \le \delta$

$$
\|u_{\Delta x}\|_{L^\infty(\Omega \times]0,T[)} \le C_T^\delta.
$$

This lemma also provides the boundedness of the discrete equilibrium states, as defined in (27).

Let (η, q) be a boundary entropy pair and let $w \in \mathbb{R}$ such that $\eta(w) = \eta'(w) = q(w) = 0$. By Lemma 3.7 there is a consistent numerical entropy flux G, so that the cell entropy inequality (39) is satisfied.

Let $\varphi \in C_0^\infty(\overline{\Omega} \times [0,T[)$, $\varphi \ge 0$. Then φ can be extended to a function in $C_0^\infty(\overline{\Omega} \times [0,T[) \cap C_0^2(\mathbb{R}^2)$, which we will also denote as φ.

Setting $C := \max\{C_T^\delta, |w|\}$ all the premises of Lemma 4.3 are fulfilled and we can deduce that for $0 < \Delta x \le \delta$ and for all non-negative test functions $\varphi \in C_0^\infty(\overline{\Omega} \times [0,T[) \cap C_0^2(\mathbb{R}^2)$, we have (45).

Now Lemma 2.4 allows us to find a sequence $(\Delta x_k)_{k \in \mathbb{N}} \subset]0, \delta]$ and a $\mu \in L^\infty(\Omega_T \times]0, 1[)$ such that for all $g \in C(\mathbb{R})$ we have

$$g\left(u_{\Delta x_k}\right) \to \int_0^1 g\left(\mu\left(\cdot, \cdot, \alpha\right)\right) \mathrm{d}\alpha \quad \text{weak} * \text{ in } L^\infty\left(\Omega_T\right). \tag{54}$$

Since $\eta, q, \eta' b \in C(\mathbb{R})$ and $\partial_t \varphi, \partial_x \varphi, z' \varphi \in L^1(\Omega_T)$ we can pass to the limit in (45) and obtain (6). Thus μ is an entropy process solution of (1) and by Theorem 2.6 get a $u \in L^\infty(\Omega_T)$ such that $\mu(\cdot, \cdot, \alpha) = u$ a.e. for almost all $\alpha \in]0, 1[$.

Since the entropy process solution is unique by Theorem 2.6, we conclude that any subsequence of $(u_{\Delta x})_{\Delta x > 0}$ converges to weak$*$ to u as $\Delta x \to 0$. By Lemma 4.4 we see that $(u_{\Delta x})_{\Delta x > 0}$ converges strongly to u in $L^p(\Omega_T)$.

5 Numerical Experiments

In this section we will numerically test the well-balanced scheme (see Definition 3.1) and compare it to the standard discretization (23). Therefore we take a look at the Burgers-Hopf equation

$$\partial_t u + \partial_x \frac{u^2}{2} + z'(x) u = 0 \tag{55}$$

on $\Omega :=]0, 4[$, which means $D(s) = s$. So (4) is obviously satisfied. For the source term we choose

$$z(x) = \begin{cases} \cos(\pi x) & \text{for } x \in \left(\frac{3}{2}, \frac{5}{2}\right), \\ 0 & \text{otherwise} \end{cases} . \tag{56}$$

Notice that z' is discontinuous and $u_c(x) := c - z(x)$ is an equilibrium of (55).

Test Case 5.1 Assume the equilibrium initial data $u_2 := 2 - z(x)$ and the constant boundary data $u_l = u_r = 2$. The exact solution to this problem is given by $u(x, t) = u_2(x)$.

Now we use 40 nodes in space and compare the numerical solutions of both schemes (as defined in (3.1) and (23) respectively) at $t = 3$. The result is shown in Figure 1.

The Tables 1 and 2 show the L^1-error for the numerical solution obtained for different grid sizes Δx. They also display the L^1-error of the numerical solution with respect to the projection of the exact solution to the space of piecewise constant functions (Numerical Error).

In these computations Δt is computed such that the CFL-condition is satisfied. The necessary estimate on $|u_j^n|$ is given by Lemma 3.5. For the standard discretisation (23) we have

$$\left|u_j^{n+1}\right| \le \sup \left|u_j^n\right| + \Delta t \left\|z'\right\|_{L^\infty(\Omega)} \sup \left|u_j^n\right| \le M e^{T\|z'\|_{L^\infty(\Omega)}} \le C_T^{\Delta x},$$

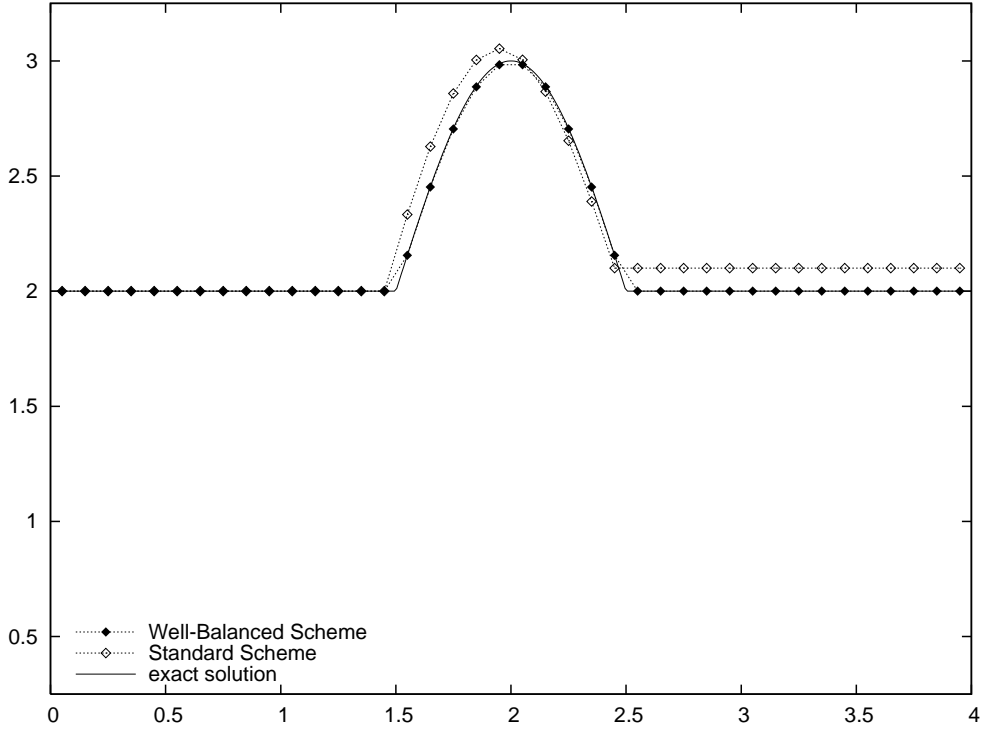where $C_T^{\Delta x}$ is defined as in (34). So our estimate for the well-balanced scheme seems far from optimal.

Figure 1: Well-Balanced and Standard Scheme for Test Case 5.1 at $t = 3$

Δx	Δt	L^1-Error	Num. Error	CPU-Time
10^{-1}	$5.7 \cdot 10^{-6}$	$5.02 \cdot 10^{-2}$	0	$0.62\,s$
10^{-2}	$3.5 \cdot 10^{-6}$	$5.00 \cdot 10^{-3}$	0	$8.83\,s$
10^{-3}	$7.3 \cdot 10^{-7}$	$5.00 \cdot 10^{-4}$	0	$423.2\,s$
10^{-4}	$8 \cdot 10^{-8}$	$5.00 \cdot 10^{-5}$	0	$42\,688\,s$

Table 1: Well-Balanced Scheme for Equilibrium Initial Data $u_0 = z\,(x)$

Δx	Δt	L^1-Error	Num. Error	CPU-Time
$2 \cdot 10^{-2}$	$4.5 \cdot 10^{-6}$	$5.07 \cdot 10^{-2}$	$4.90 \cdot 10^{-2}$	$2.94\,s$
$2 \cdot 10^{-3}$	$1.3 \cdot 10^{-6}$	$5.09 \cdot 10^{-3}$	$4.91 \cdot 10^{-3}$	$97.6\,s$
$2 \cdot 10^{-4}$	$1.6 \cdot 10^{-7}$	$5.09 \cdot 10^{-4}$	$4.92 \cdot 10^{-4}$	$10\,885\,s$
$2 \cdot 10^{-5}$	$2 \cdot 10^{-8}$	$5.06 \cdot 10^{-5}$	$4.97 \cdot 10^{-5}$	—

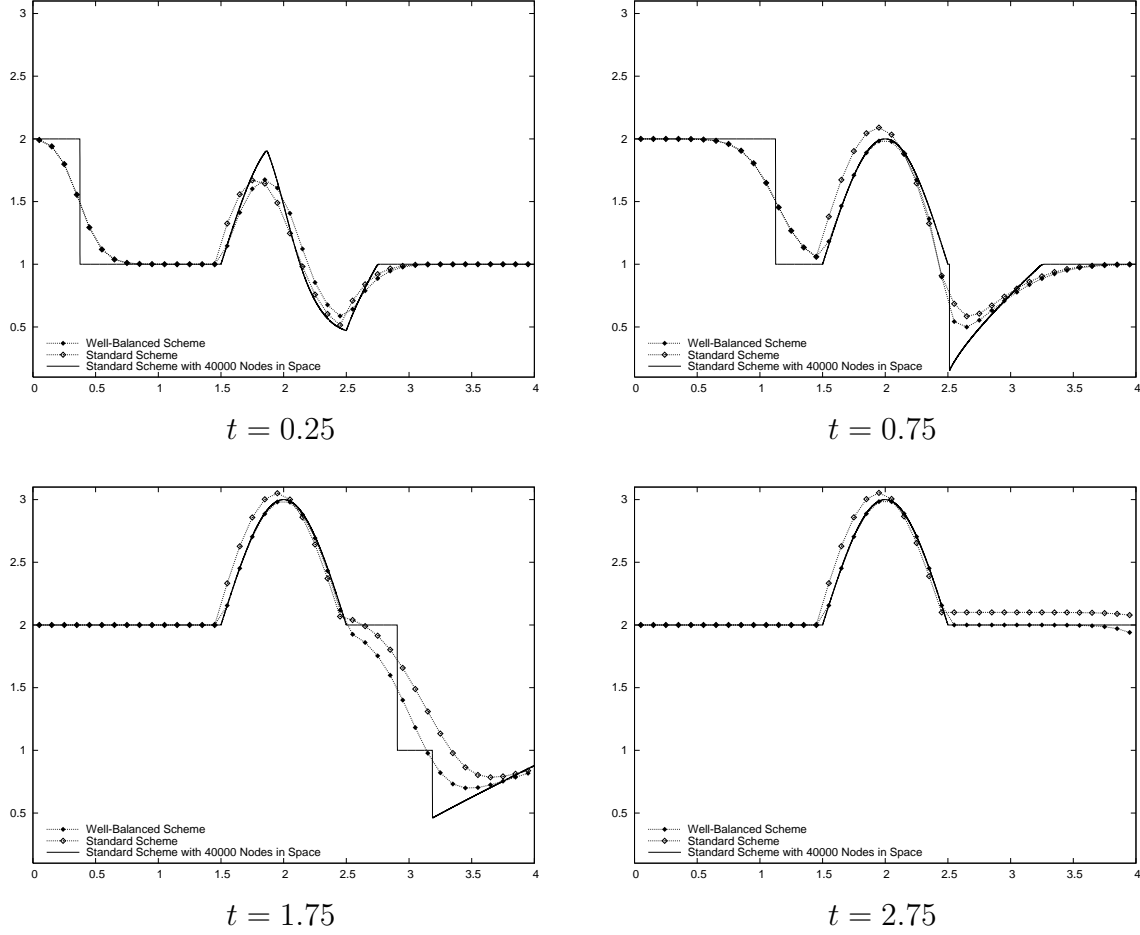Table 2: Standard Scheme for Equilibrium Initial Data $2 - z\,(x)$

23

Figure 2: Well-balanced scheme vs. standard discretization for Test Case 5.2

Test Case 5.2 Assume the constant initial data $u_0 = 1$ and the constant boundary data $u_l = 2$, $u_r = 1$. Again we use 40 nodes in space to compare the well-balanced scheme as defined in (3.1) to the standard discretization (23). A reference solution was obtained by the standard scheme with $40\,000$ nodes in space. The result is displayed in Figure 5.2.

Now we turn to the weakness of the well-balanced scheme. For the Burgers-Hopf equation (55) $u(x,t) = 0$ is an equilibrium for any choice of z. Thus, if we ask the boundary data $u_l = u_r = 0$ there are two equilibria satisfying this condition: $-z(x)$ and 0.

Test Case 5.3 We assume the constant initial data $u_0 = 0$ and the constant boundary data $u_l = u_r = 0$. the exact solution is given by $u(x,t) = 0$. Figure 3 displays the numerical solutions obtained by the well-balanced scheme for $40, 400, 4\,000$ and $40\,000$ nodes in space at $t = 2.5$.

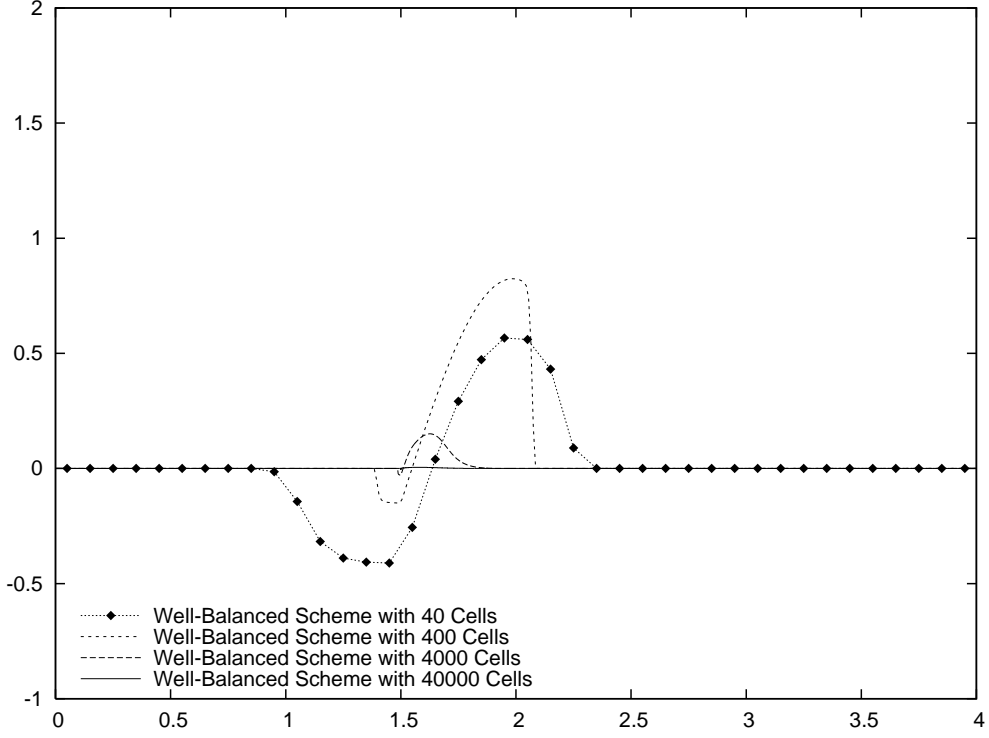Table 3 shows the L^1-error for this problem, which is not acceptable even for $4\,000$ nodes in space.

Figure 3: Well-balanced scheme for Test Case 5.3 at $t = 2.5$

Δx	Δt	L^1-Error	CPU-Time
10^{-1}	$6.14 \cdot 10^{-6}$	$4.388 \cdot 10^{-1}$	$0.90\,s$
10^{-2}	$6.14 \cdot 10^{-6}$	$3.164 \cdot 10^{-1}$	$5.23\,s$
10^{-3}	$6.14 \cdot 10^{-6}$	$2.678 \cdot 10^{-2}$	$47.9\,s$
10^{-4}	$6.14 \cdot 10^{-6}$	$8.421 \cdot 10^{-4}$	$740.8\,s$

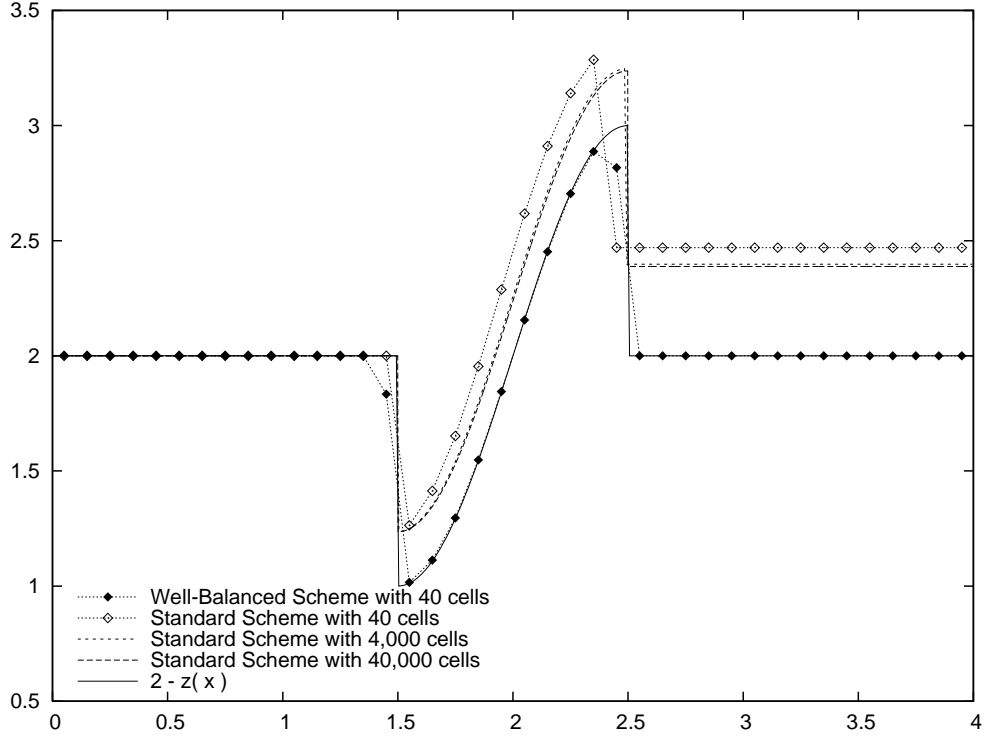Table 3: Well-Balanced Scheme for Initial Data $u_0 = 0$

Figure 4: Well-Balanced and Standard Scheme for Test Case 5.4 at $t = 3$

Notice that the convergence proof for the well-balanced scheme holds for $u = 0$, even though we divided the differential equation in (1) by u in the motivation.

We also tested both the well-balanced scheme and the standard scheme (23) for a discontinuous z:

Test Case 5.4 Let z be given by the discontiuous function

$$z(x) = \begin{cases} \sin(\pi x) & \text{for } x \in \left(\frac{3}{2}, \frac{5}{2}\right), \\ 0 & otherwise \end{cases}$$

and choose the initial data $u_2 = 2 - z(x)$ and the constant boundary data $u_l = u_r = 2$. The numerical results in this case are displayed in Figure 4.

Note that while the well-balanced scheme converges to the equilibrium quite fast, the standard scheme seems not to converge to the equilibrium at all.

References

[1] Bardos, C., Le Roux, A.Y., Nedelec, J.C.: First Order Quasilinear Equations with Boundary Conditions, Communications in Partial Differential Equations **4**, 1017–1034 (1979)

[2] Botchorishvili, R.: Equilibrium Type Schemes for Scalar Conservation Laws with Source Term, Applied Mathematics and Informatics **7**, 1–34 (2002)

[3] Botchorishvili, R., Perthame, B., Vasseur, A.: Equilibrium Schemes for Scalar Conservation Laws with Stiff Sources, Mathematics of Computation **72**, 131–157 (2003)

[4] R. Botchorisshivili, O. Pironneau: *Finite Volume Schems with Equilibrium Type Discretization of Source Terms for Scalar Conservation Laws*, Journal of Computational Physics 187, No. 2, 2003, 391–427.

[5] Brenier, Y.: Résolution d'équations d'évolution quasilinéaires en dimension N d'espace à l'aide d'équations linéaires en dimensions $N + 1$, Journal of Differential Equations **50**, 375–390 (1982)

[6] Engquist, B., Osher, S.: One-Sided Difference Approximations for Nonlinear Conservation Laws, Mathematics of Computation **36**, 321–351 (1981)

[7] Eymard, R., Gallouët, T., Herbin, R.: Existence and Uniqueness of the Entopy Solution to a Nonlinear Hyperbolic Equation, Chinese Annals of Mathematics, Series B **16**, 1–14 (1995)

[8] Gosse, L.: A Priori Error Estimate for a Well-Balanced Scheme Designed for Inhomogenous Scalar Conservation Laws, Comptes Rendus de l'Académie des Science, Série I: Mathématique **327**, 467–472 (1998)

[9] Greenberg, J.M., LeRoux, A.Y.: A Well-Balanced Scheme for the Numerical Processing of Source Terms in Hyperbolic Equations, SIAM Journal of Numerical Analysis **33**, 1–16 (1996)

[10] Kröner, D.: *Numerical Schemes for Conservation Laws* (Wiley-Teubner, 1997)

[11] Kružkov, S.N.: First Order Quasilinear Equations in Several Independent Variables, Mathematics of the USSR: Sbornik **10**, 217 – 243 (1970)

[12] Málek, J., Nečas, J., Rokyta, M., Ružička, M.: Weak and Measure-Valued Solutions to Evolutionary PDEs (Chapman & Hall, 1996)

[13] Nolte, M.: Ein balanciertes Verfahren zur numerischen Lösung von Anfangsrandwertproblemen für skalare Erhaltungsgleichungen mit Quellterm in 1D

[14] Otto, F.: Initial-Boundary Value Problem for a Scalar Conservation Law, Comptes Rendus de l'Académie des Science, Série I: Mathématique **322**, 729–734 (1996)

[15] Vovelle, J.: Convergence of Finite Volume Monotone Schemes for Scalar Conservation Laws on Bounded Domains, Numerische Mathematik **90**, 563–596 (2002)